POSTPROCESSED TIMESCALES AT THE U.S. NAVAL OBSERVATORY

Demetrios Matsakis and Lee A. Breakiron
U.S. Naval Observatory
Washington, DC 20392, USA

Abstract

A set of free-running timescales is generated using 9 years of data from the USNO clock ensemble. Cesium and maser clock phases and frequencies are characterized by global fits to first- or second-order polynomials, and the timescales are generated from the clocks' detrended frequencies using a variety of weighting functions. These timescales are compared to those generated by USNO's A.1, BIPM's ALGOS and TT98, and NIST's AT1 algorithms from the perspective of algorithm choice and frequency stability.

1 THE ALGORITHM

Since its derivation and implementation by Don Percival, the U.S. Naval Observatory (USNO)'s mean algorithm[1] has proven robust and adaptable for over 20 years. It is the basis of the USNO free-running mean timescale A.1, which is used to form the USNO mean timescale that is frequency-steered to International Atomic Time (TAI) or, equivalently, UTC(BIPM). In turn, the USNO Master Clock (MC2) is steered to the USNO mean timescale to generate UTC(USNO), providing the most precise on-line realization of UTC in the world today.

The A.1 timescale is actually an integrated frequency scale. In its current formulation, clock frequencies (rates) are generated by averaging the hourly first differences of clock timing data referenced to the USNO Master Clock (MC2). Individual clock frequencies are detrended, using the past A.1 as a reference, through a first-order fit (effectively, to clock rate and drift) over time ranges of uniform clock behavior, as determined by experienced USNO data analysts. Clocks showing poor performance, or not yet well-determined rates and drifts, are ignored, while all others are included with a weight depending upon clock type and, for masers, the time since the present[2]. For example, this dynamic weighting system initially weights day-old data from cavity-tuned masers up to 12 times more than data from cesium clocks, and 3-day-old maser data 5 times more than contemporaneous cesium data. It completely deweights maser data older than 75 days.

The A.1 timescale was always intended to be a compromise between stability and robustness. Its algorithm was motivated by the need to optimize on-line precision for a clock ensemble with a limited number of masers; different algorithms are now under consideration for the 12-maser ensemble currently maintained at the USNO. One innovation introduced in January, 1999 is to detrend cesiums and masers against a reference timescale composed of only cesium standards, as opposed to A.1. Another innovation under test is to realize UTC through a timescale composed of only masers that have been detrended against the unsteered cesium average.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **DEC 1998**	2. REPORT TYPE	3. DATES COVERED **00-00-1998 to 00-00-1998**

4. TITLE AND SUBTITLE **Postprocessed Timescales at the U.S. Naval Observatory**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Naval Observatory,3450 Massachusetts Ave NW,Washington,DC,20392**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR’S ACRONYM(S)
	11. SPONSOR/MONITOR’S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADA415578. 30th Annual Precise Time and Time Interval (PTTI) Systems and Applications Meeting, Reston, VA, 1-3 Dec 1998

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**15**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Another possible problem with the current A.1 algorithm is that clocks are detrended individually, instead of all at once. The differences between the global and local minima for the values of the clock characterization polynomials are not very important in an on-line timescale whose main purpose is to provide a template for monthly steers to TAI; however, they are more important for work relating to comparisons with other free-running time scales or pulsar data.

We report here timescales generated using a postprocessed algorithm, informally titled SuperP, whose detrending polynomial models are determined through a global fit to inter-clock phase difference data which has been differenced from the temporally preceding clock difference N times. For example, phase data differenced once ("first differences") are equivalent to frequency data. Note that because the clock data are recorded only in the form of differences between clocks, the choice of reference is irrelevant. Also, the SuperP and A.1 timescales are underdetermined, by a polynomial of order M, if the timescale is generated from N integrations of a scale based upon Nth-order differences, which are detrended using a polynomial of order M-N. For solutions involving drift-corrections (M=2), any parabola may be added to the final timescale without affecting the consistency of the solution for optimal polynomial detrending coefficients. The free-running timescales TA(NIST), generated by NIST using their AT1 algorithm, and EAL, generated by the BIPM using their algorithm ALGOS[3], are also sensitive to the initial timescale reference (Table 1); if the initial values of one of these timescales had differed by an constant and slope, that same difference would have persisted to the present. If a perfectly calibrated set of drift differences between external timescales were available, it would be possible to determine the parabolic term from a limited set. One consequence of using all the data to resolve the indeterminancy, as opposed to a subset that is assumed to be better calibrated, is that long-term variations, such as would be expected due to white frequency noise, are masked and the effective errors in the comparisons are increased[4].

2 THE DATA

The USNO maintains an on-line archive of (currently) 9 years of clock data from its maser and cesium ensemble, beginning on MJD 47752 (11 August, 1989) and ending on MJD 51086 (30 September, 1998). Although lower-noise measurement systems are also being used, this work is based only on data taken with a time-interval counter and switch system, whose measurement accuracy is better than 100 ps. For brevity, only measurements at 0 hours UT were used in this analysis. In Figure 1, the numbers of each type of clock producing acceptable data are presented. Unfortunately, the decisions made by data analysts for maser data previous to MJD 50079 were not permanently recorded; thus, maser data previous to that time are here ignored, although re-analysis may be made at a later date. It is also possible that the editing information available for the oldest data is not accurate, and that a re-analysis will improve the results slightly. As is evident in the later figures, timescales are of lesser stability previous to about MJD 49400 (17 February, 1994); the subsequent improvement is due to the dramatic increase in the number and quality of clocks maintained at the USNO and contributing to TAI (cavity-tuned Sigma-Tau/Datum masers and HP5071 cesiums). The natural division of the data into three time ranges is the reason why different analyses presented below will begin at different MJD's. Comparisons with TT98 are limited by its cutoff in December, 1997.

This work is based on only data recorded as acceptable by USNO data analysts, with some additional automatic editing of outliers identified through a simple median-based scheme. The time ranges of the on-line polynomial clock characterizations determined by the data analysts were used to define the break points in the global least-squares solution to the polynomials. The important question of whether these time ranges are optimal in the determination of rate

and drifts is not addressed here. The A.1 used here differs somewhat from what has been reported to the BIPM due to the effect of dynamic weighting and occasional refinements in editing or clock characterization made after submission to the BIPM.

For comparisons with non-USNO timescales, time-transfer noise has always been a serious problem. Although much better than previous modes of time transfer, even common-view GPS time transfer has been shown to have systematic errors on the order of tens of nanoseconds. In 1995 and 1996 a BIPM-organized calibration effort revealed a total 29-ns calibration error in the common-view chain between USNO, NIST, and the Observatory of Paris (OP), and somewhat lower errors in the links between many other institutions and OP. Although most of the institutions involved immediately had their data adjusted with a single time-step, USNO's data were gradually corrected, in 3-ns steps, over the year 1997 (MJD 50482-50783), in consideration of USNO's users, most of whom require greater frequency stability. It is possible to verify the results of this procedure in a rough manner, by comparing the values for GPS time reported in Circular T with those measured at the USNO (Figure 2). Although this comparison does not benefit from the common-view removal of Selective Availability (SA), the 29-ns change is apparent, along with what appears to be a small (4 ns) residual error. This small error may be due to remaining calibration errors anywhere in the measurement chain. In the comparisons with UTC, EAL, and TT98, it was found that correcting the differences as if this problem had never occurred did not improve the comparisons, nor was there any improvement after crude allowance was made for the high weight USNO clocks have in the generation of EAL. It is possible that improvement would be evident if adjustments could be made for the fact that many other institutions had time-transfer corrections made at the same time.

3 THE SuperP FAMILY OF TIMESCALES

Perhaps the most important of the controversial issues related to timescale algorithms is the determination of clock weights, which need not be the same in the clock characterization and the timescale generation, and at times could be zero.

Although clocks whose frequency data display Gaussian statistics should theoretically be weighted by the inverse variance of their frequency data, in practice the USNO has found it more robust to weight all clocks of the same type equally[5]. This approach is justified because the accuracy and precision of measurement systems and clocks are difficult to assess, partly due to nonzero covariances in clock performance data[6,7], and because clock time series are not statistically stationary; in particular, past performance is not always a good indicator of future performance.

To study the effect of weighting, a very incomplete set of 9 functions was explored for weighting individual differences between clocks in the clock characterization solution. In all cases, measured difference data for each pair of clocks were weighted as the inverse root-sum-square (RSS) of the set of clock weights given in Table 2.

In applying the weights to the clock characterization, all possible clock pairs at each MJD were used without allowance for the strong correlation between pairs that include the same clock. Although solutions could be generated restricting all pairs to those involving any single reference clock (which may change between measurement times), this would complicate any solutions involving a variance-based weighting scheme. A more robust approach would involve using the full covariance function in determining a non-diagonal weight matrix for clock characterization. Of these weighting schemes, the best one (based on the 3-cornered-hat analysis described in the next section), effectively removes masers from the characterization process by down-weighting them by a factor of 1000.

Once the clocks were detrended using the clock characterization determined by the fitting, timescales were generated as averages of all clocks, one point per day, using the same weighting schemes as for clock characterization. For clarity in the analysis, the problem of combining data from masers with cesiums to form a timescale was bypassed through the generation of separate pure-maser and pure-cesium timescales. We note that modern steering theory would provide an optimal way to steer a supposedly less-accurate, but more stable, maser-based timescale to a cesium-based one[8].

4 THE INTERNAL ERRORS

The different weighting options were examined through the consistency of timescales generated using independent subsets of 1/3 of the data, and performing a 3-cornered-hat analysis which allowed for possible data correlations using a technique that minimizes covariances [7]. The subsets were chosen by assigning clocks in the order they were encountered by the computer, and excluded data before MJD 49400. The resulting stability assessments for cesium-only and maser-only averages (Figures 3) indicate a weak preference for a weighting scheme in which maser and cesium clocks are characterized by comparison with a unity-weighted pure-cesium mean, but the improvement of ignoring masers for clock characterization should be larger in an on-line timescale generated using our current dynamic weighting scheme.

Using the SuperP formalism, it is simple to generate timescales from other than the data's first differences, and to compare their results. Through determination of the internal errors, using the same technique as above, it was found that fitting first-order polynomials to the first differences (frequencies, as is done for A.1, AT1, and EAL) was preferable to fitting second-order polynomials to phase data, constants to second difference data, or nothing to third difference data (Figures 4). This was also found using the A.1 formalism[9], and is expected in a situation dominated by white FM. Once the clock characterization has been determined, differencing to order N also has the effect of smoothing over phase discontinuities of order N-1 that may be associated with gaps in the data. Since all the free-running timescales considered here are generated from first differences, this result validates what has long been done in practice.

Figures 5 show how the 3-cornered-hat analysis estimates the stability of HP5071-only and Sigma-Tau/Datum-only timescales derived by integrating average detrended frequencies, for which the clock characterization was determined through a weighting scheme sensitive only to cesiums and down-weighting non-HP5071 cesiums by a factor of 0.65. Also shown are a curve for the mean timescale of one-half of the clocks relative to the mean timescale of the other half, where the deviations were reduced in size by a factor of $\sqrt{2}$ to convert them from relative error to absolute error (neglecting covariances), and a curve for the mean of the entire HP5071 ensemble (assuming the three subsets could be weighted according to their inverse variances, hence also neglecting covariances).

5 THE EXTERNAL ERRORS

To estimate the external errors in the USNO data, comparisons were made between the USNO timescales and the BIPM timescales EAL and TT98 (the latter is in essence a postprocessed UTC, determined from EAL using information available at the end of 1997[10]), and their frequency stabilities relative to TT98 are shown in Figure 6. The most stable timescales are those generated by SuperP using only the most recent data; however, the disparity between

the size of the clock ensembles utilized and the existence of unmodelled time transfer noise obscures these comparisons.

Figures 7-11 compare the A.1, SuperP pure-maser average, SuperP pure-cesium average, and TA(NIST) with EAL and TT98. Here, A.1 is essentially a pure-cesium timescale because all the maser data have been phased out except for the last 75 days by the dynamic weighting. Again, for recent data the smallest RMS errors relative to TT98 were found for the SuperP timescales. For comparisons going further back into the past, A.1 provides the best fit to EAL. It is difficult to draw firm conclusions from these comparisons because it is evident that the statistics are not stationary. It would be better to simply note that the comparisons reflect the considerable improvement in recent years. This improvement is also evident in the USNO's ability to better steer our Master Clock to UTC—as reported in BIPM's last Circular T, all the timing differences between the USNO Master Clock and UTC were less than 2 ns. While we don't expect this close alignment to continue in the near future, it is entirely possible that pending improvements in time transfer and frequency standards may result in such differences between UTC and UTC(USNO) becoming routine in several more years.

6 CONCLUSIONS

A set of free-running timescales were generated using 9 years of USNO clock data and an algorithm dubbed SuperP, which made global fits of phases and frequencies to first-order and second-order polynomials. USNO clock frequencies and drifts are currently determined with respect to a pure cesium-based timescale. The procedure proved to be the best of those treated herein.

Timescales were generated from the clocks' detrended frequencies and a variety of weighting functions. Frequency stability assessments indicated a preference for fitting first-order polynomials to first differences, rather than other polynomials or other types of data.

Comparisons of A.1, SuperP, and TA(NIST) timescales to TT98 showed the greatest frequency stability for those of SuperP, while the A.1 provided the best fit to EAL. Partial allowance for past USNO/BIPM calibration errors does not improve the comparison between USNO data and the BIPM-generated timescales.

7 ACKNOWLEDGMENTS

We thank Randy Clarke and Harold Chadsey for their years of labor examining individual clock data, and the entire staff of the USNO Time Service Department for maintaining and improving the Master Clock and its many components, including those needed for time transfer to the international timekeeping community.

8 REFERENCES

[1] D. B. Percival 1978, *"The U.S. Naval Observatory clock time scales,"* **IEEE Trans. Instr. Meas., IM-27**, 376-385.

[2] L. A. Breakiron 1992, *"Timescale algorithms combining cesium clocks and hydrogen masers,"* Proceedings of the 23rd Annual Precise Time and Time Interval (PTTI)

Applications and Planning Meeting, 3-5 December, 1991, Pasadena, California, USA, pp. 297-305.

[3] C. Thomas, P. Wolf, and P. Tavella 1994, **Time Scales**, Monograph 94/1, Bureau International des Poids et Mesures, Sèvres, France.

[4] J. A. Barnes 1984, *"The measurement of linear frequency drift in oscillators,"* Proceedings of the 15th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 6-8 December 1983, Washington, DC, USA, pp. 551-582.

[5] L. A. Breakiron 1995, *"The effects of clock errors on timescale stability,"* Proceedings of the 26th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 6-8 December, 1994, Reston, Virginia, USA, pp. 369-380 (see Errata volume #2).

[6] A. Premoli and P. Tavella 1993, *"A revisited three-cornered hat method for estimating frequency standard instability,"* **IEEE Trans. Instr. Meas., IM-42**, 7-13.

[7] F. Torcaso, C. R. Ekstrom, E. A. Burt, and D.N. Matsakis 1999, *"Estimating frequency stability and cross-correlations,"* these Proceedings.

[8] D. Matsakis, M. Miranian, and P. Koppang 1999, *"Steering the U.S. Naval Observatory (USNO) Master Clock,"* Proceedings of 1999 Institute of Navigation Technical Meeting, 25-27 January 1999, San Diego, California, USA, in press.

[9] L. A. Breakiron 1994, *"A comparative study of clock rate and drift estimation,"* Proceedings of the 25th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting, 29 November-2 December 1993, Marina del Rey, California, USA, pp. 401-412.

[10] B. Guinot 1988, *"Atomic time scales for pulsar studies and other demanding applications,"* **Astronomy and Astrophysics, 192**, 370-373.

Table 1. Comparison of several algorithms with SuperP.

Algorithm	Timescale Weights	Constant Correction?	Freq Corr?	Drift Corr?	Detrending Time (days)
NIST (AT1)	(exp)*	N/A	yes	(exp)*	20-30
EAL(BIPM)**	(1-year var)	N/A	yes	no	30 (was 60)
UTC,TT(BIPM)	***	N/A	***	0***	-
USNO (A.1)	dynamic	N/A	yes	yes	as needed****
SuperP					
Ave phase	8 options	option	option	option	as needed
Ave 1st diff	"	N/A	option	option	"
Ave 2nd diff	"	N/A	N/A	option	"
Ave 3rd diff	"	N/A	N/A	N/A	"

* NIST AT1 based on exponential filter: last frequency estimate averaged with most recent estimate, with time constant set to 20-30 days. Weights are based upon the inverse variance and a similar exponential filter[3].

** BIPM-EAL subtracts from each clock phase a term: A+B*time, where A is the previous phase and B is the frequency estimate. This would be 100% equivalent to second diff if their "B" were related to frequency obtained from two adjacent 5-day points instead of the past 30 days (formerly 60). Clocks with significant drifts are deweighted[3].

*** UTC and TT are generated by steering UTC to the primary (calibrated) frequency standards, hence the drift is zero by definition. Of course time-transfer noise and frequency measurement errors are not completely negligible.

**** Typically 30-360 days.

Table 2. Clock-based weight schemes explored in this work. Schemes 4 through 9 are based upon the performance (statistical properties) of the clock, as measured through an initial computation of the difference between the detrended clock data and an average of all clocks, using unity weights to characterize and average the initial estimate, which is performed in the same difference mode as the final computations.

1. 1.0 for all clocks
2. 1.0 for HP5071 cesiums, .65 for other cesiums, .001 for masers
3. 1.0 for HP5071 cesiums and all masers, .65 for other cesiums
4. Inverse of squared sum of temporarily adjacent points
5. Inverse of variance, computed after removal of mean difference
6. Inverse variance
7. Inverse variance times the time range of the data in MJD
8. Inverse variance times the time-range of the data squared
9. Inverse variance times the time-range of the data cubed

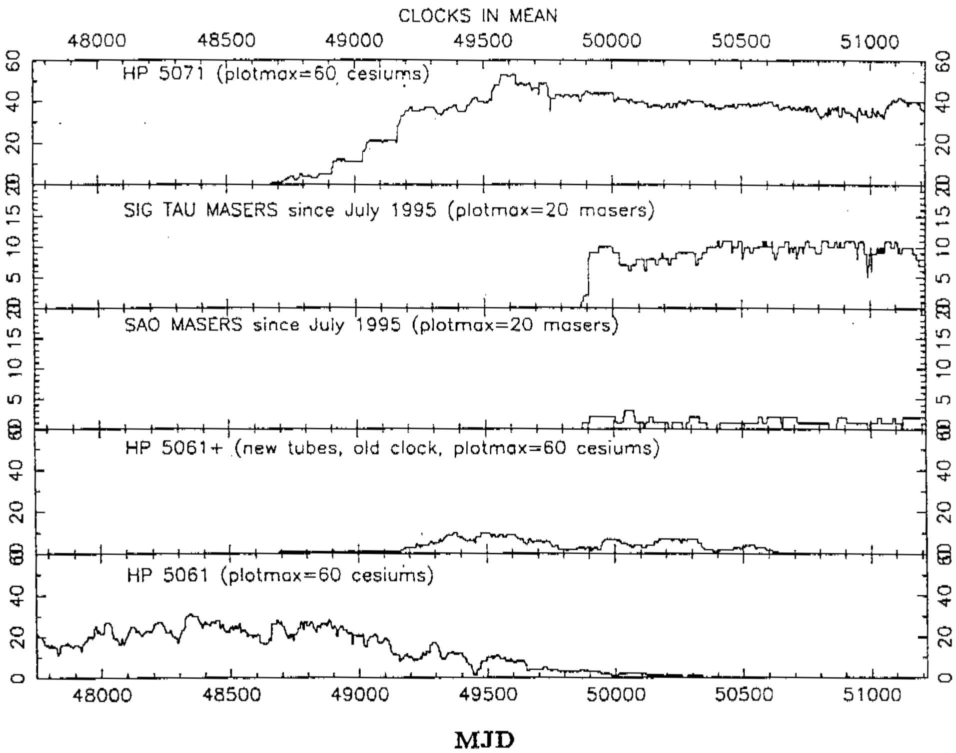

Figure 1. Number of USNO clocks used in SuperP and A.1. Maser data previous to MJD 50079 are available, but the associated editing and trend-break decisions would have to be re-evaluated.

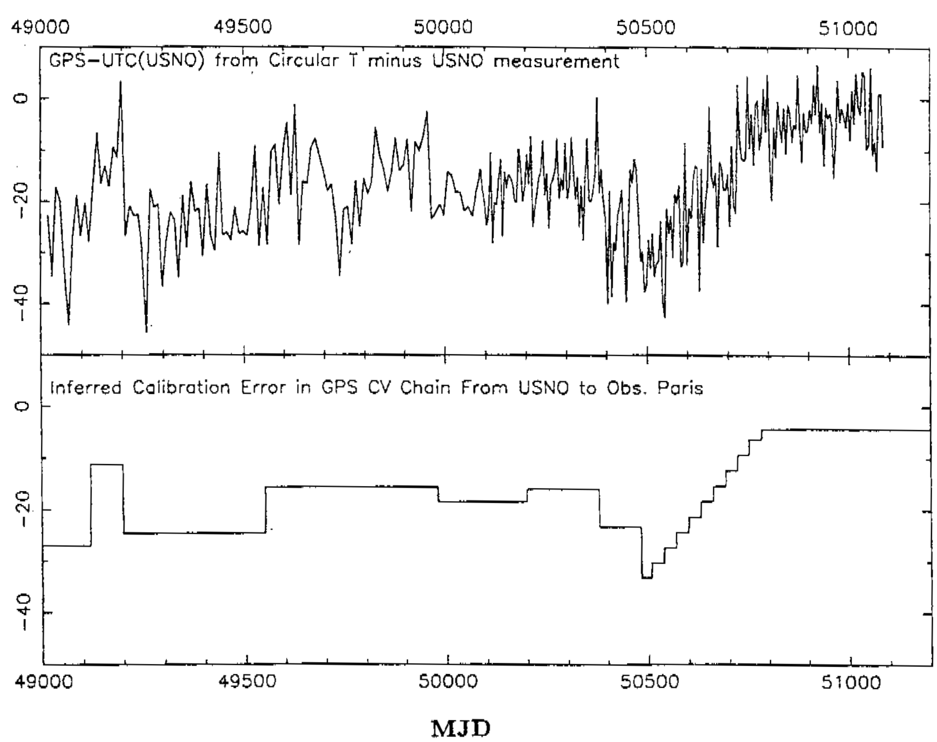

Figure 2. Calibration errors in chain USNO-NIST-OP, as estimated from USNO and BIPM measurements of GPS.

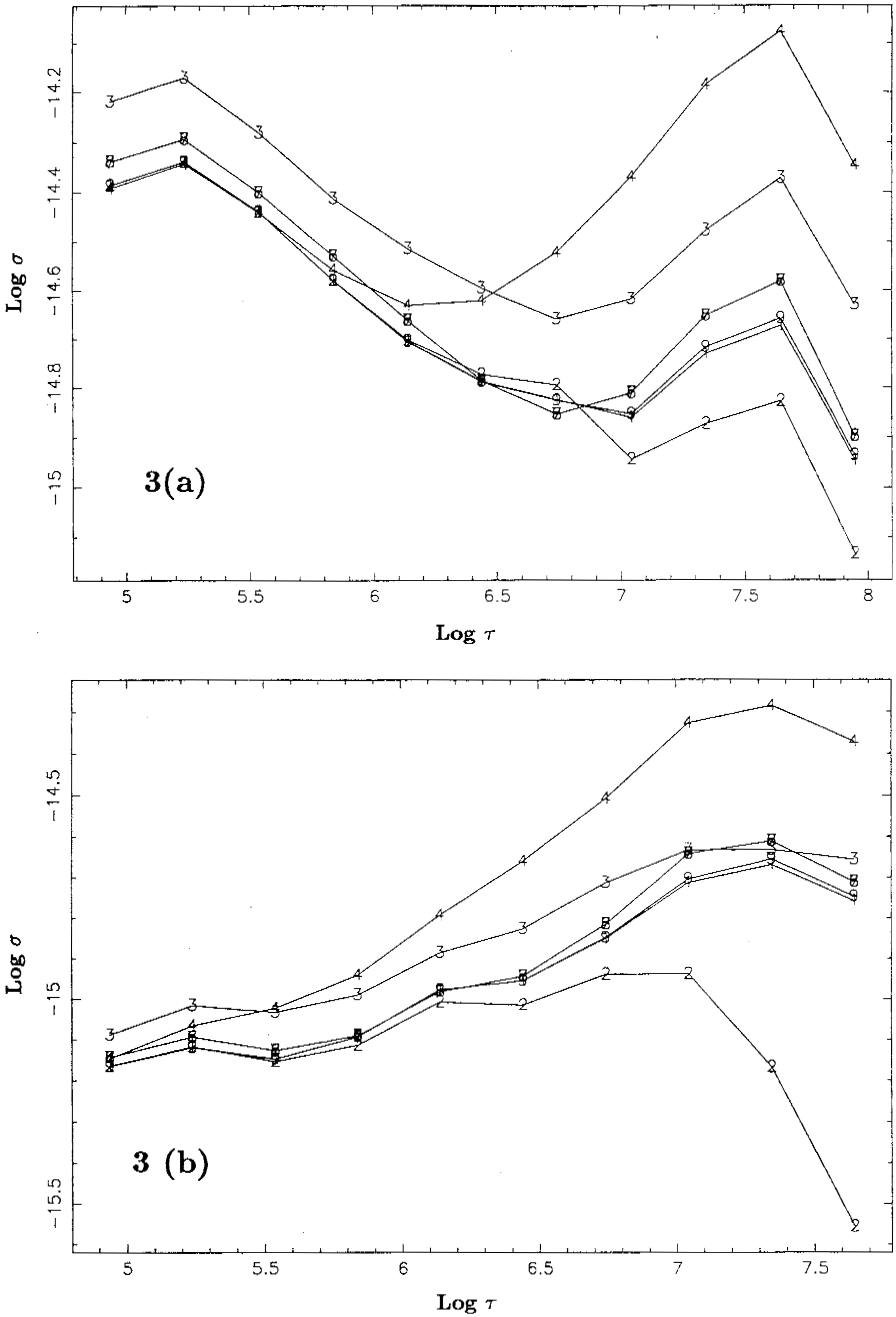

Figure 3. Frequency stabilities for integrated first-difference timescales derived using weighting schemes in Table 2. (a) is for cesium-only data since MJD 49400. (b) is for maser-only data since MJD 50079. Plot characters correspond to rows of Table 2. σ is the Allan deviation for frequency and τ is the sampling time in seconds.

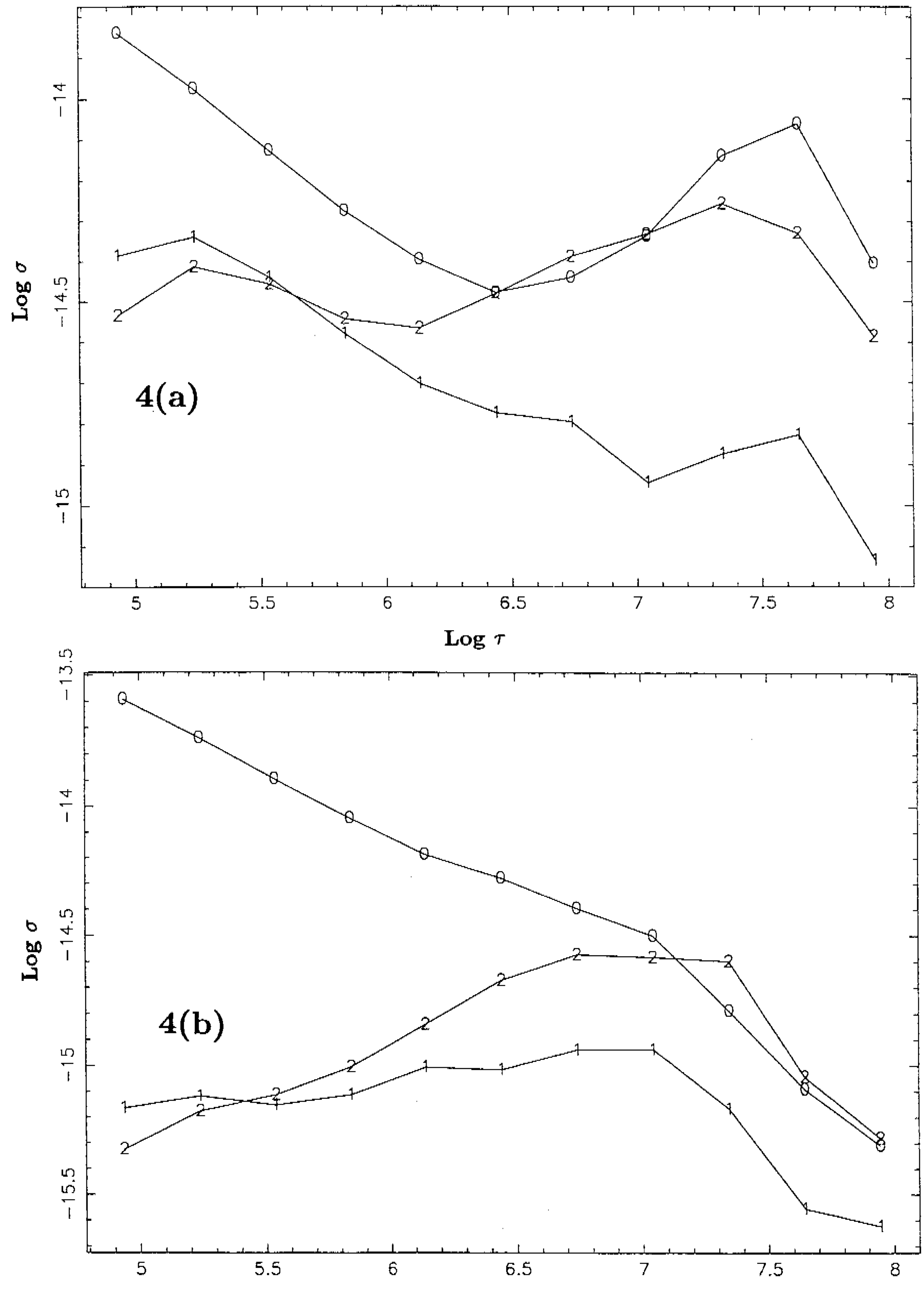

Figure 4. Frequency stabilities for timescales derived using temporal Nth-order data differences and cesium-dominated weight scheme "2." (a) is for cesium-only timescales using data since MJD 49400. (b) is for maser-only timescales derived using data since MJD 50079. Plot characters correspond to order of pre-fit differencing.

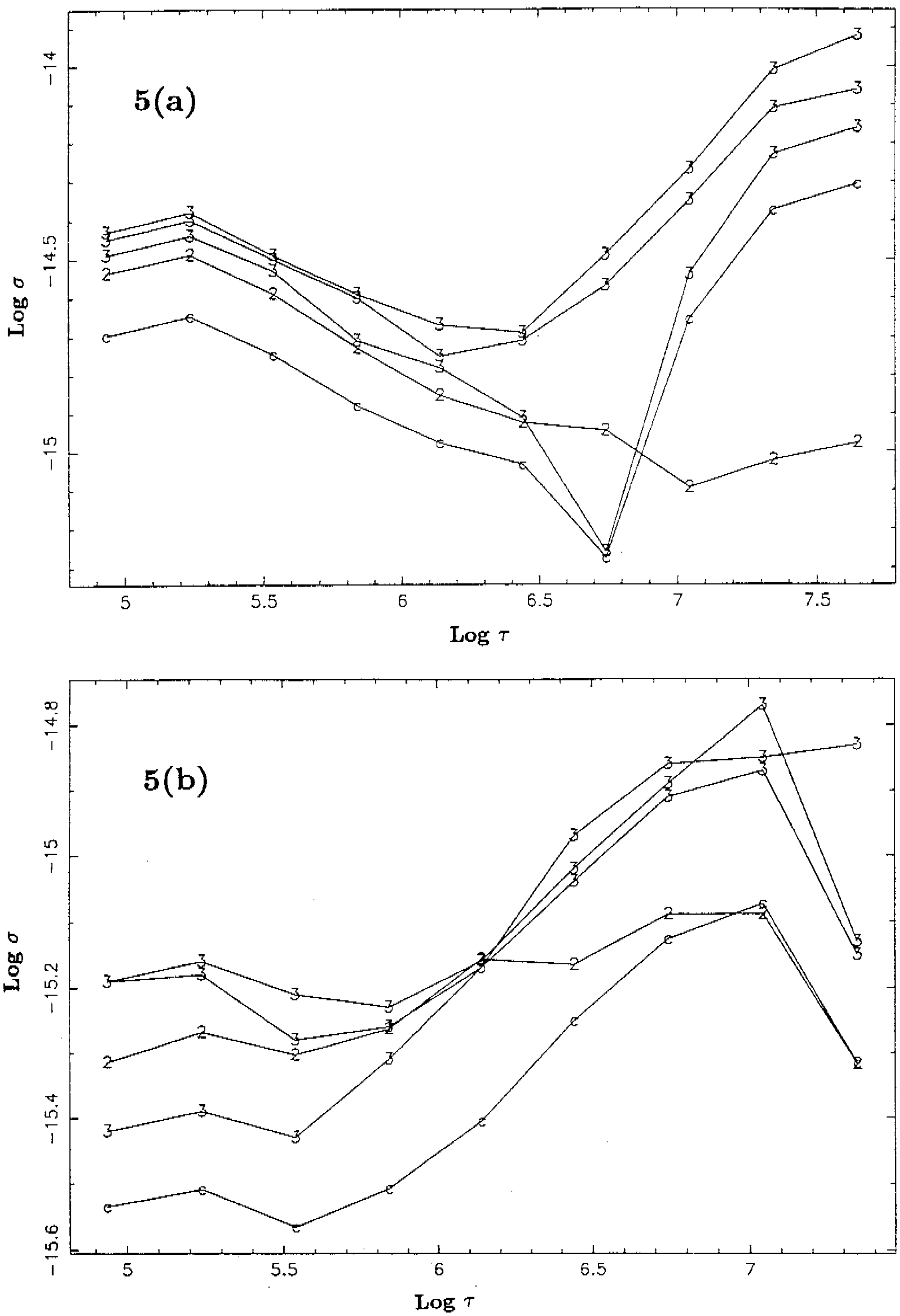

Figure 5. Frequency stabilities for mean timescales generated using only one type of clock and three independent subsets, derived from a 3-cornered-hat analysis of timescales that allowed for covariances. (a) is for an HP5071-only timescale and is based upon data since MJD 49400. (b) is for a Sigma Tau/Datum maser-only timescale and is based on data since MJD 50079. The curve indicated by "2" is for a mean timescale of one-half of the clocks relative to a mean timescale of the other half, and the curve indicated by "c" is for the complete USNO HP5071 ensemble (in both, covariances were neglected).

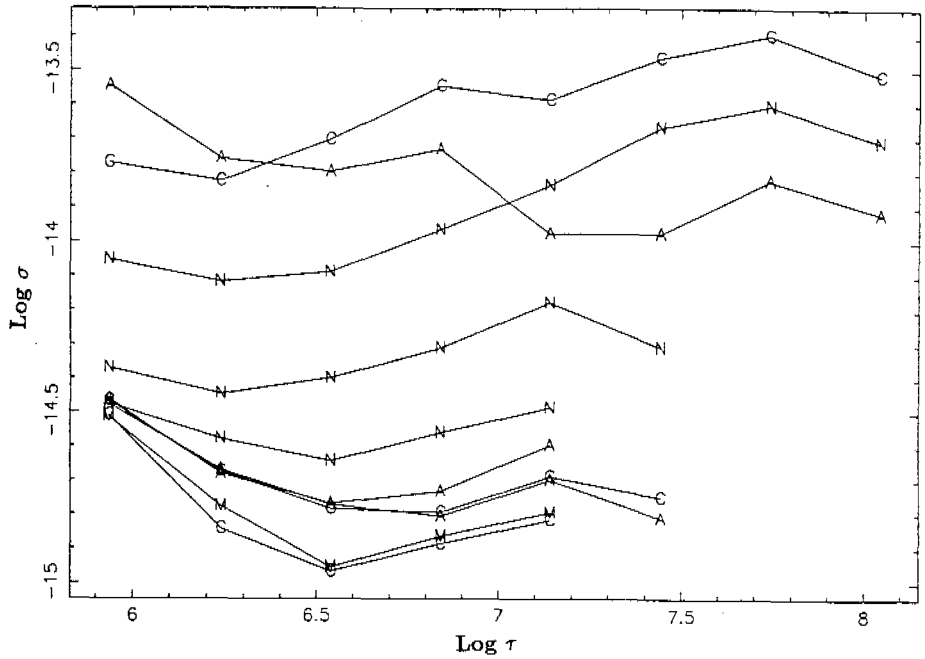

Figure 6. Frequency stability plots comparing timescales to TT98 over three different sampling times τ. Those of TT98-A.1 are denoted **A**, TT98-SuperP cesium-only average are denoted C, TT98-SuperP maser-only averages are denoted M, and TT98-TA(NIST) are denoted N. The three curves that extend over the longest range of log τ include all data since MJD 44752; the cesium-only average is the only such to include non-HP5071 cesiums. The three curves of shortest range in log τ are for data since MJD 50079, and the three of intermediate range are for data since MJD 49400.

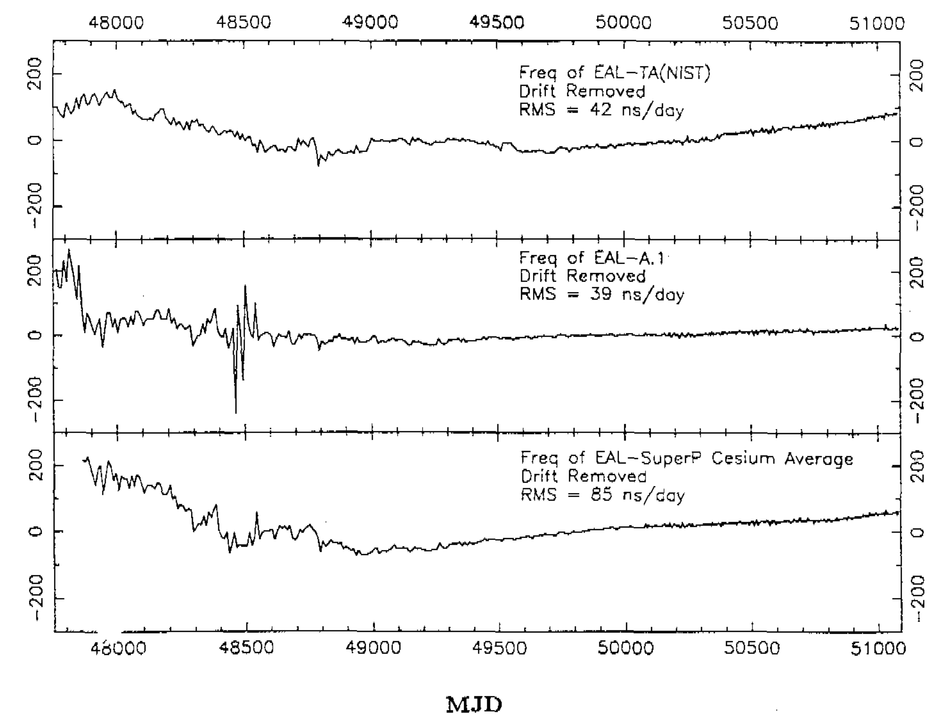

Figure 7. Frequencies in ns/day of EAL-TA(NIST), EAL-A.1, and EAL-SuperP cesium-only average. For display purposes, some high-frequency data from the SuperP timescale were removed previous to MJD 47900; however, they are still included in the RMS calculation.

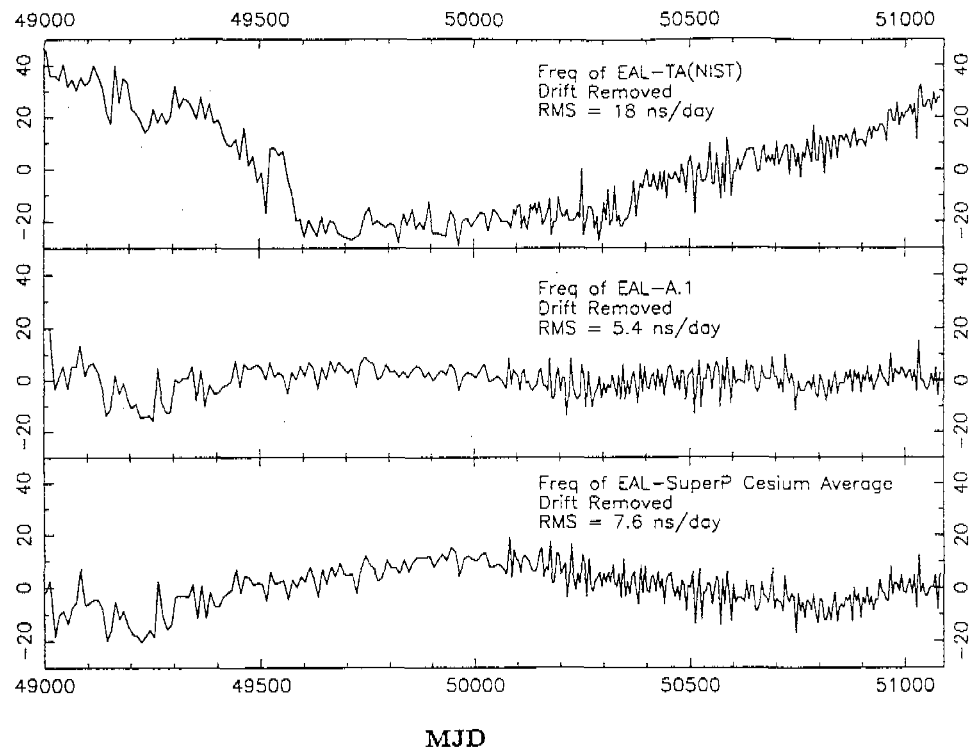

Figure 8. Frequencies in ns/day of EAL-TA(NIST), EAL-A.1, and EAL-SuperP using only data since MJD 49000.

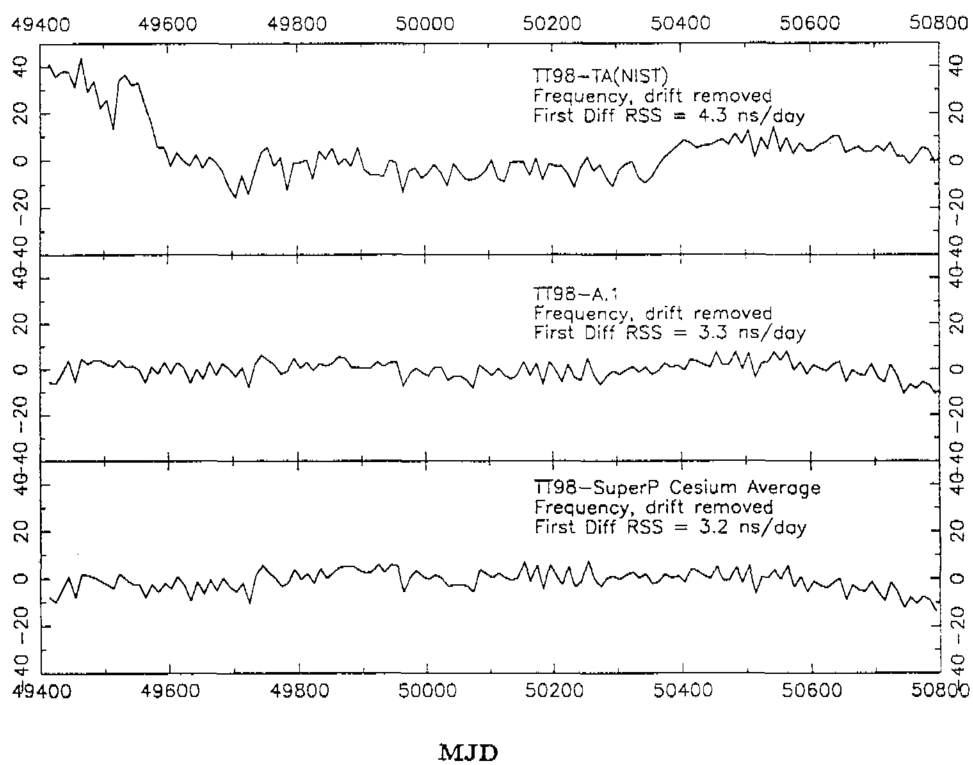

Figure 9. Frequencies in ns/day of TT98-TA(NIST), TT98-A.1, and TT98-SuperP using only data since MJD 49400.

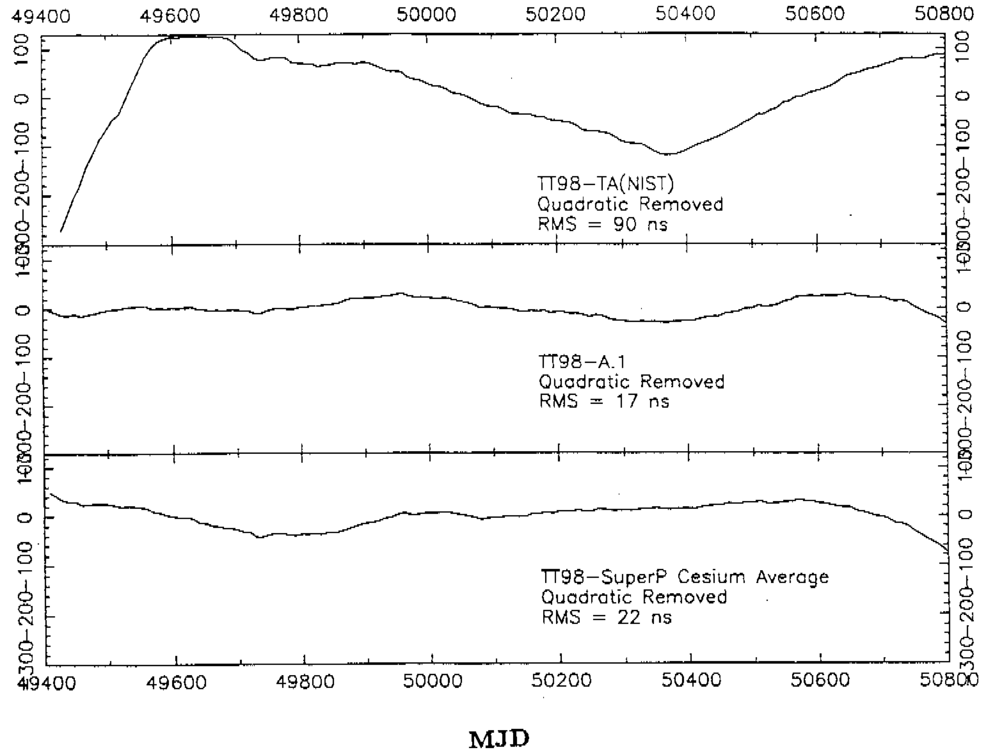

Figure 10. Differences in ns between timescales and TT98 using only data since MJD 49400.

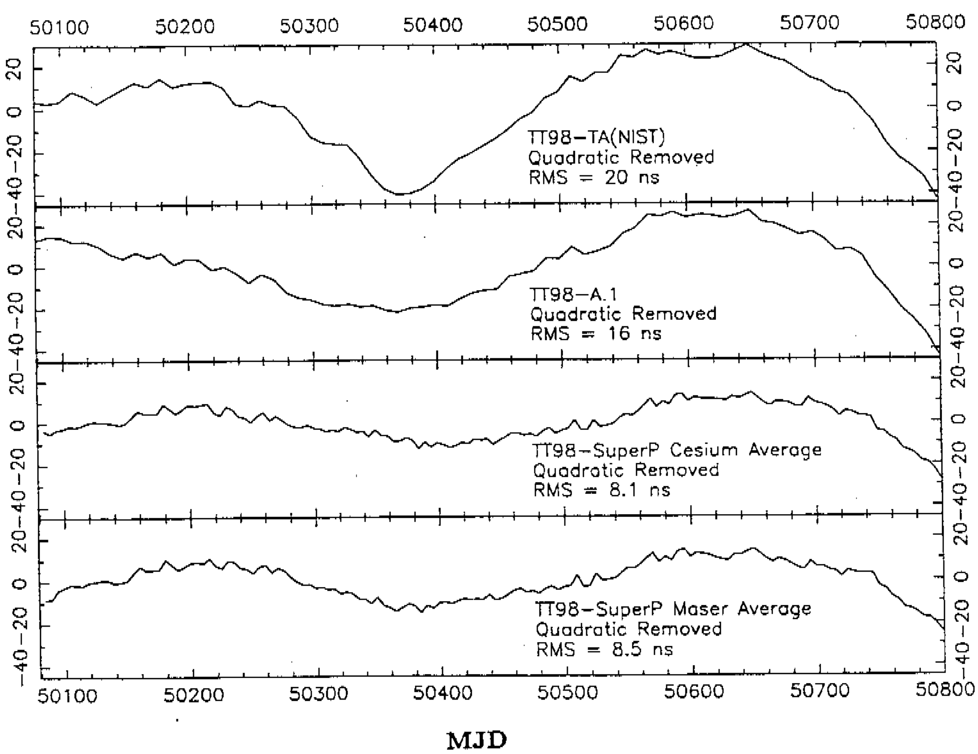

Figure 11. Differences in ns between timescales and TT98 using only data since MJD 50079.

Questions and Answers

GERARD PETIT (BIPM): Your comparison with TT-BIPM is, of course, comparing two different kinds of scales. TT-BIPM is one algorithm designed to provide the longest possible time span, one scale based upon existing primary standards. It provides something, which has been done consistently for 20 years and your plot is over 20 years. Maybe one confusion is that data from the primary standards are not sufficient compared to the quality of the clocks in the past two years.

DEMETRIOS MATSAKIS (USNO): I think that is absolutely the case. I think we have to look to the future, not to the past. Certainly those kinds of errors which consider 2-nanosecond variations 10 or 15 years ago would be impossible; you would not even think about it. Everything was much worse back then; but now everything is getting better.

MARC WEISS (NIST): I have a question: In the removal of a quadratic in order to compare time scales, I would be concerned that first of all, when you remove a quadratic, you remove a fair amount of the random walk from the scale; so it is difficult to see the difference between a quadratic and a random walk over a short period of time, even over years. Secondly, the drift of the clocks comes in and produces a quadratic, and it seems like modeling the drift is a big problem. I think it is problematic to remove a drift, that is, to remove a third of a quadratic from the time scale in order to compare them.

DEMETRIOS MATSAKIS: Well, it is in there. I could change my original estimate of the drift 10 years ago, and it would show right now. So you can keep it in and pretend it is not there, but you would only be fooling yourself. You could compare frequencies instead of time because this is really a frequency scale which is integrated. You would get the same results. I mean the same general conclusions.